**Principal Component Analysis**

**Introduction:**

Principal Component Analysis (PCA) is a statistical technique used for dimensionality reduction, data compression, and feature extraction. It transforms a large set of variables into a smaller one that still contains most of the information in the large set. This method is widely used in fields like machine learning, data analysis, and bioinformatics.

**Objectives:**

1. Dimensionality Reduction: Reduce the number of variables while preserving as much information as possible.
2. Data Visualization: Visualize high-dimensional data in 2D or 3D
3. Noise Reduction: Eliminate noise and redundant features from the dataset.
4. Feature Extraction: Identify and derive new, meaningful features from the original dataset.

**Some Mathematical operations PCA involved in:**

Standardize the Data, Compute the Covariance Matrix, Calculate Eigenvalues and Eigenvectors, Sort Eigenvalues and Eigenvectors, Project the Data

**Applications:**

1. **Image Compression**: Reduce the dimensionality of image data while preserving key features.
2. **Finance**: Identify the most influential factors in stock prices.
3. **Genomics**: Analyze gene expression data to find significant patterns.
4. **Marketing**: Segment customers based on purchasing behavior.

**Advantages:**

* **Simplicity**: Easy to implement and interpret.
* **Efficiency**: Reduces computational cost by reducing dimensions.
* **Noise Reduction**: Helps in denoising the data.

**Disadvantages:**

* **Linearity**: Assumes linear relationships between variables.
* **Interpretability**: Principal components may not have a straightforward interpretation.
* **Information Loss**: Some information may be lost during dimensionality reduction.

**Resources:** <https://builtin.com> , <https://www.nature.com/articles/nmeth> , <https://www.analyticsvidhya.com> , <https://www.geeksforgeeks.org> , <https://chatgpt.com>